



# Hotel Booking Analysis

Project - EDA-Hotels

## Project Overview

### ► Business Context

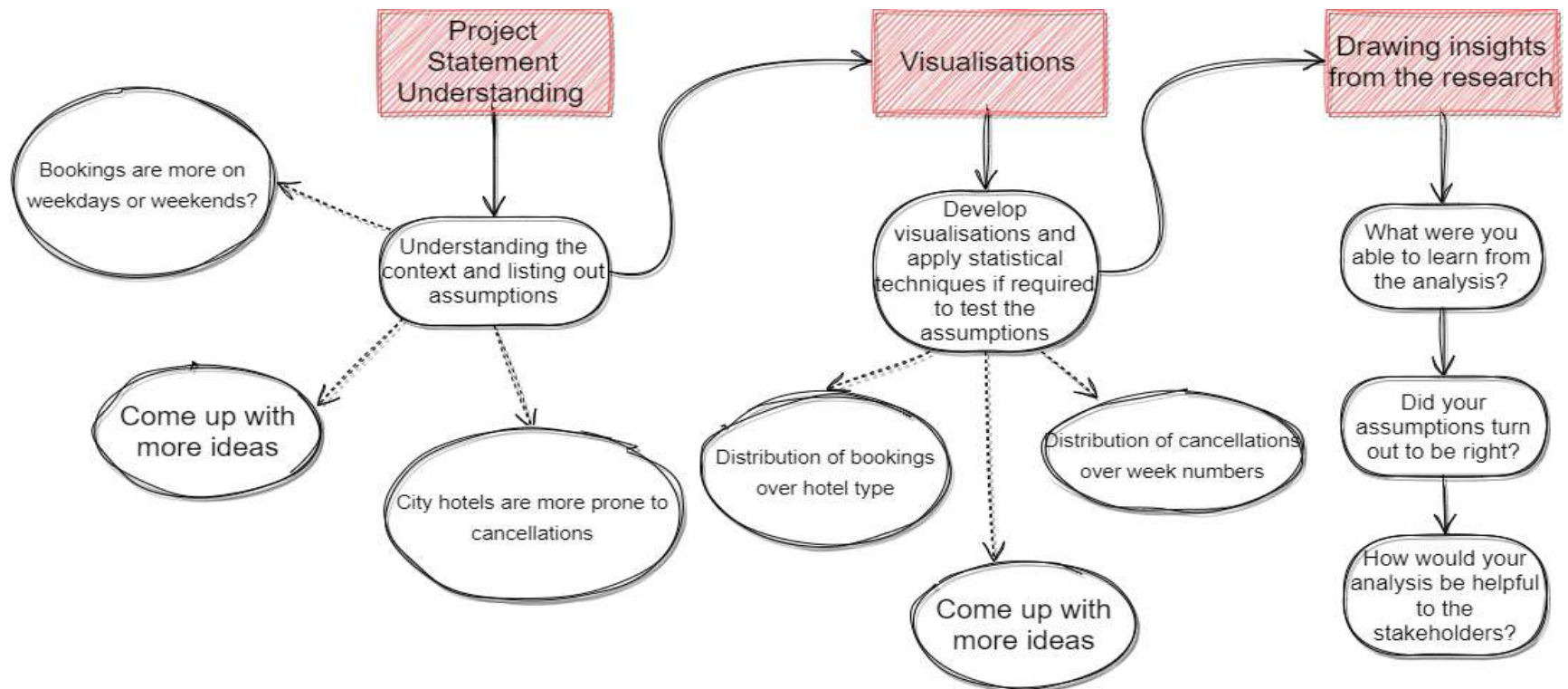
Have you ever wondered when the best time of year to book a hotel room is? Or the optimal length of stay in order to get the best daily rate? What if you wanted to predict whether or not a hotel was likely to receive a disproportionately high number of special requests? This hotel booking project can help you explore those questions!

This project contains booking information for a city hotel and a resort hotel, and includes information such as when the booking was made, length of stay, the number of adults, children, and/or babies, and the number of available parking spaces, among other things. All personally identifying information has been removed from the data. Explore and analyze the data to discover important factors that govern the bookings.

# Dataset description

Field	Description		
Hotel	H1= Resort Hotel H2=City Hotel	Distribution _channel	How the customer accessed the stay-corporate booking/Direct/TA.TO
is_cancelled	If the booking was cancelled(1) or not(0)	is_repeated_guest	Guest coming for first time or not
lead_time	Number of days that elapsed between the entering date of the booking into the PMS and the arrival date	previous_cancellation	Was there a cancellation before
arrival_date_year	Year of arrival date	previous_bookings	Count of previous bookings
arrival_date_month	Month of arrival date	reserved_room_type	Type of room reserved
arrival_date_week_number	Week number for arrival date	assigned_room_type	Type of room assigned
arrival_dat_day	Day of arrival date	booking_changes	Count of changes made to booking
stays_in_weekend_nights	Number of weekend nights (Saturday or Sunday) the guest stayed or booked to stay at the hotel	deposit_type	Deposit type
stays_in_week_nights	Number of week nights (Monday to Friday) the guest stayed or booked to stay at the hotel	agent	Booked through agent
adults	Number of adults	days_in_waiting_list	Number of days in waiting list
children	Number of children	customer_type	Type of customer
babies	Number of babies	required_car_parking	If car parking is required
meal	Kind of meal opted for	total_of_special_req	Number of additional special requirements
country	Country code	reservation_status	Reservation of status
market_segment	Which segment the customer belongs to	reservation_status_date	Date of the specific status

# Project Architecture





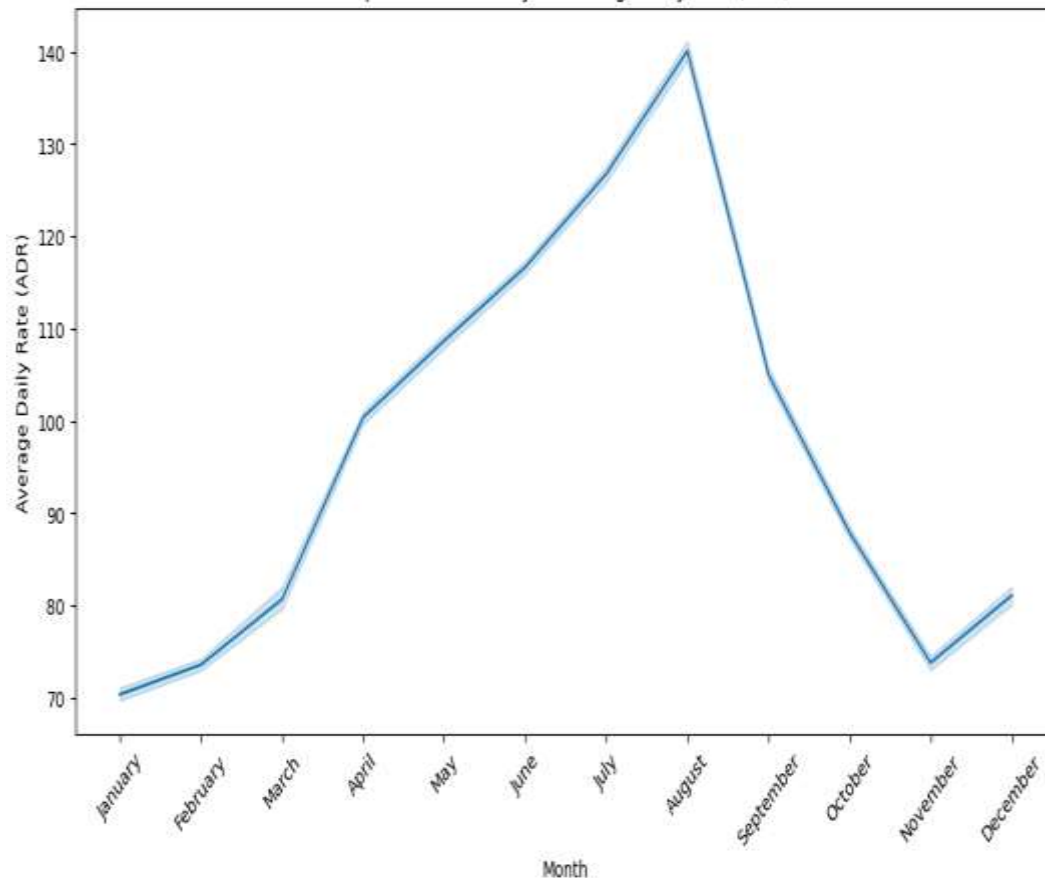
# Problem statement

- ▶ **Optimizing Pricing Strategies:** By analyzing booking patterns, you can identify peak seasons and times of low demand. This information can help hotels optimize their pricing strategies, offering competitive rates during off-peak times and maximizing revenue during high-demand periods.
- ▶ **Enhancing Customer Experience:** Understanding factors such as the length of stay, special requests, and preferences of guests can contribute to a better overall customer experience. Hotels can use this information to tailor their services and amenities to meet the specific needs and expectations of their guests.
- ▶ **Improving Marketing Efforts:** Analysis of booking data can reveal demographic information about the guests, helping hotels target specific customer segments with personalized marketing campaigns. This can lead to more effective advertising, promotions, and customer engagement strategies.
- ▶ **Forecasting Demand:** By examining historical booking data, hotels can forecast future demand more accurately. This enables better resource planning, ensuring that hotels are adequately staffed and prepared to meet the needs of guests during peak periods.
- ▶ **Optimizing Inventory Management:** Hotels can use booking data to optimize room inventory management. This involves adjusting room availability based on historical booking trends, minimizing overbooking or underutilization of rooms, and maximizing revenue per available room.
- ▶ **Enhancing Operational Efficiency:** By understanding the factors influencing the booking process, hotels can streamline their operations. This includes optimizing check-in/check-out processes, allocating resources efficiently, and ensuring that staff is well-prepared for busy periods.
- ▶ **Reducing Operational Costs:** Improved forecasting and efficiency can lead to cost savings. For example, hotels can adjust staffing levels based on expected demand, reducing unnecessary labor costs during slower periods.
- ▶ **Risk Management:** Analyzing data can help identify potential risks and challenges, allowing hotels to implement proactive measures. This includes preparing for high-demand events, managing cancellations effectively, and mitigating the impact of unforeseen circumstances.
- ▶ **Competitive Advantage:** Hotels that leverage data analytics for booking analysis gain a competitive advantage. They can stay ahead of market trends, adapt quickly to changes, and continuously improve their services based on customer feedback and preferences.

**Optimizing Pricing Strategies :** This includes various sub-problems to address the impact on pricing.

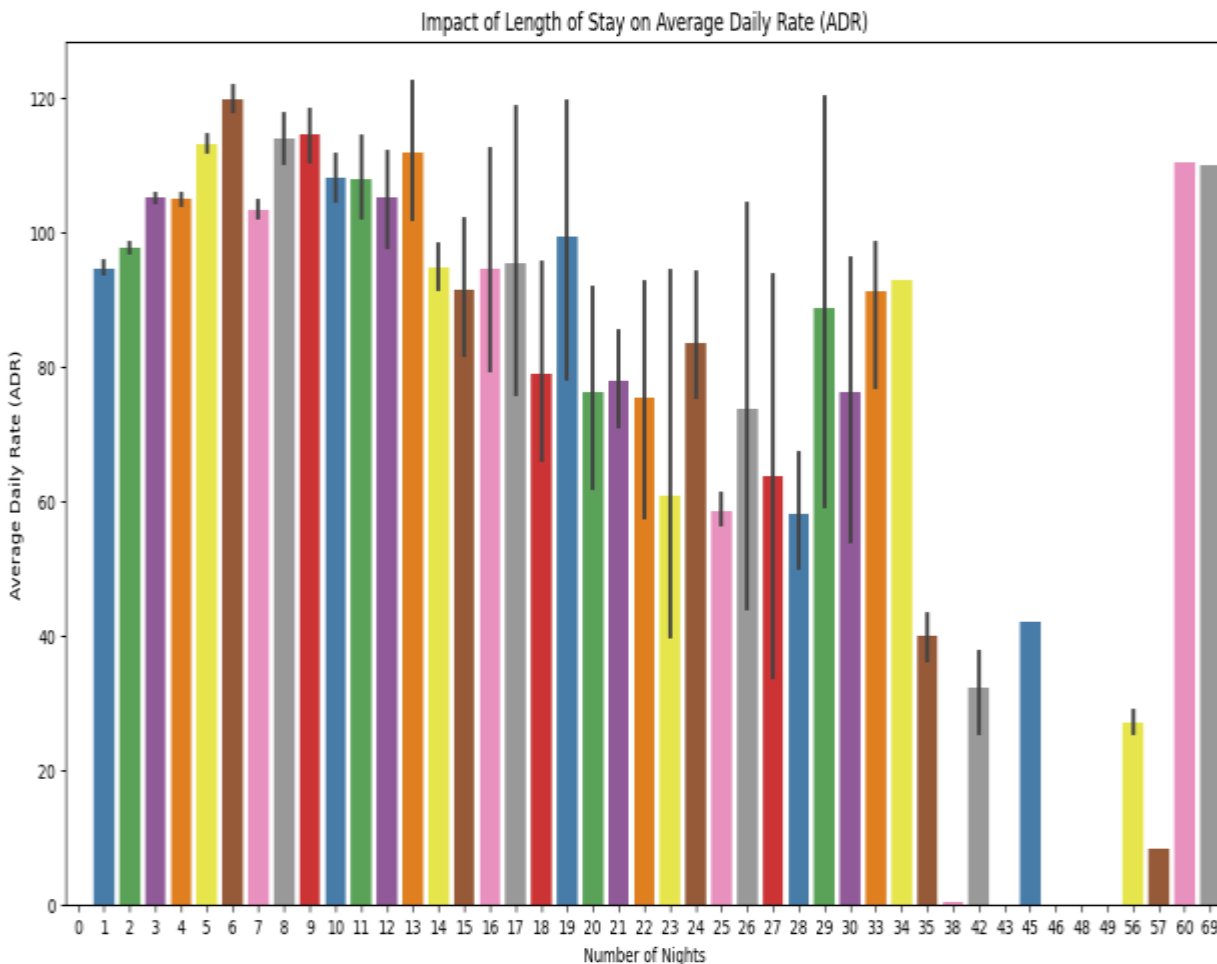
➡ **First problem as shown below**

Impact of Seasonality on Average Daily Rate (ADR)



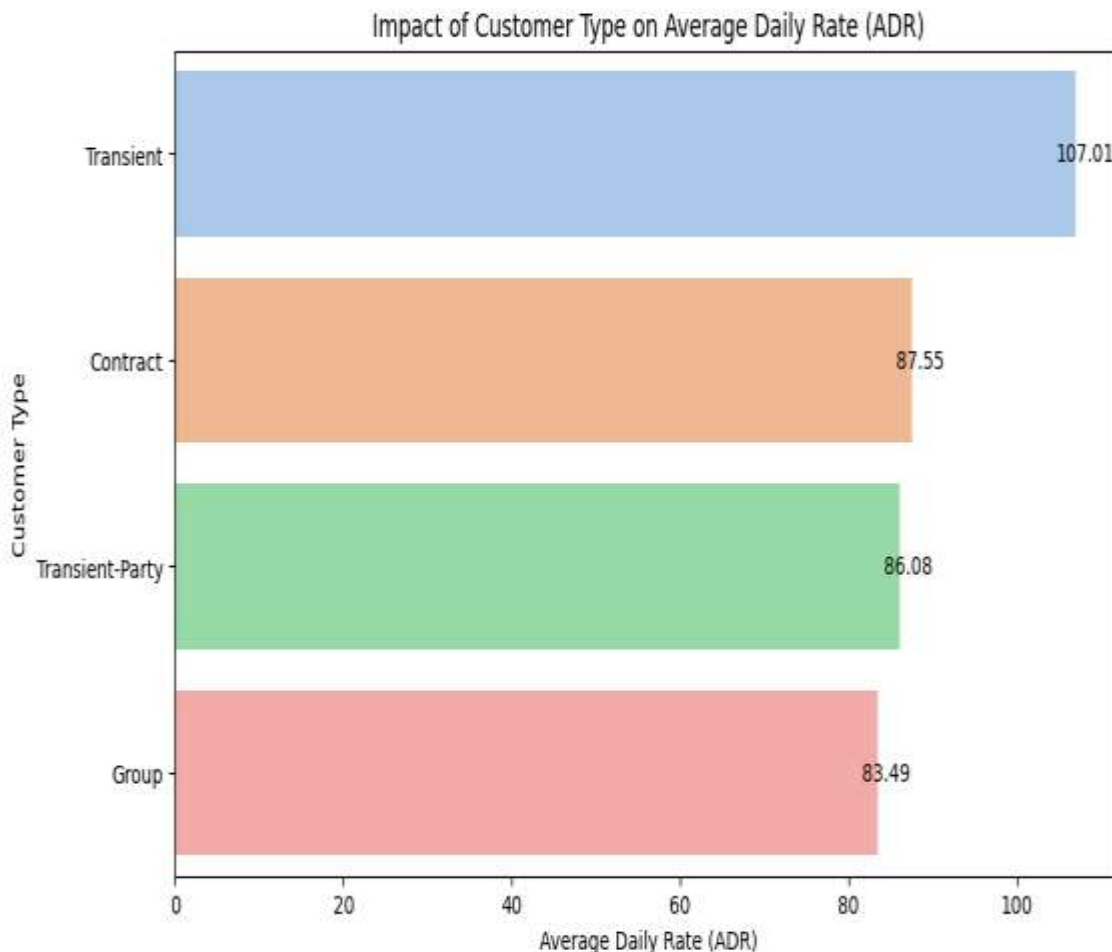
- ▶ **Seasonality Impact on Pricing:** A line chart showing the average daily rate (adr) across different months
- ▶ **Observation:** The average daily rate (adr) is peaked for the month of July and August

➡ Second problem as shown below



- ▶ **Length of Stay Impact on Pricing:** A bar chart illustrating how the length of stay influences the average daily rate (adr).
- ▶ **Observation:** The chart illustrates a relatively stable Average Daily Rate (ADR) with a noticeable increase for shorter stays, peaking within a specific range of around 2 to 15 nights and suggesting that factors other than the duration of the stay may influence pricing dynamics.

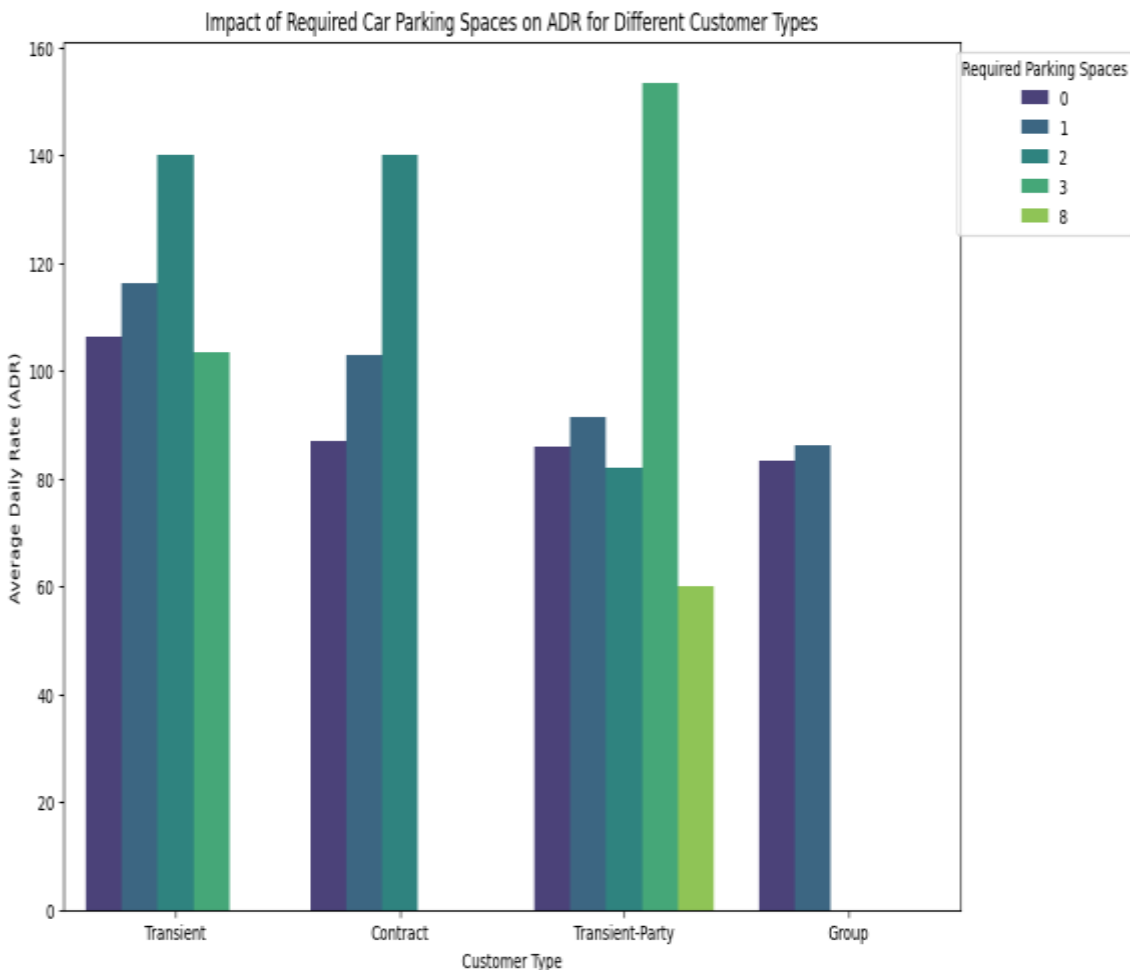
➡ Third problem as shown below



- **Impact of customer type on adr:** A bar plot showing the average daily rate (adr) across different customer type.
- **Observation:** The average daily rate (adr) is high(107.01) for Transient type of customers (typically include tourists, business travelers, or any guests making individual reservations)



➡ Forth problem as shown below



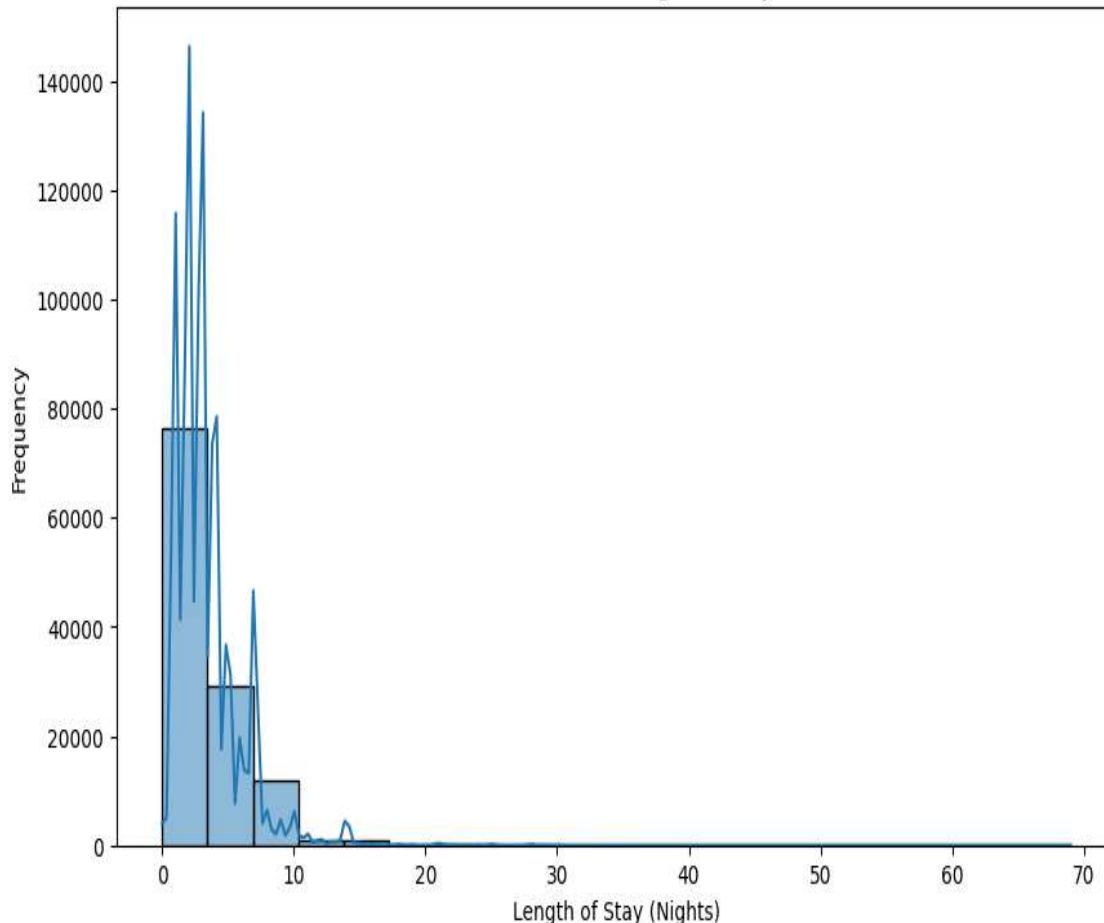
➤ **Impact of car parking space on ADR for different type of customers**

➤ **Observation:** As we have previously observed the Number of nights are less for Transient-party compared to Transient Type of customers but in this plot it is observed that Transient-party is having high requirement of car parking space which is maximizing the revenue by giving good adr.

Enhancing Customer Experience: This includes various sub-problems to address the overall customer experience.

➡ First problem as shown below

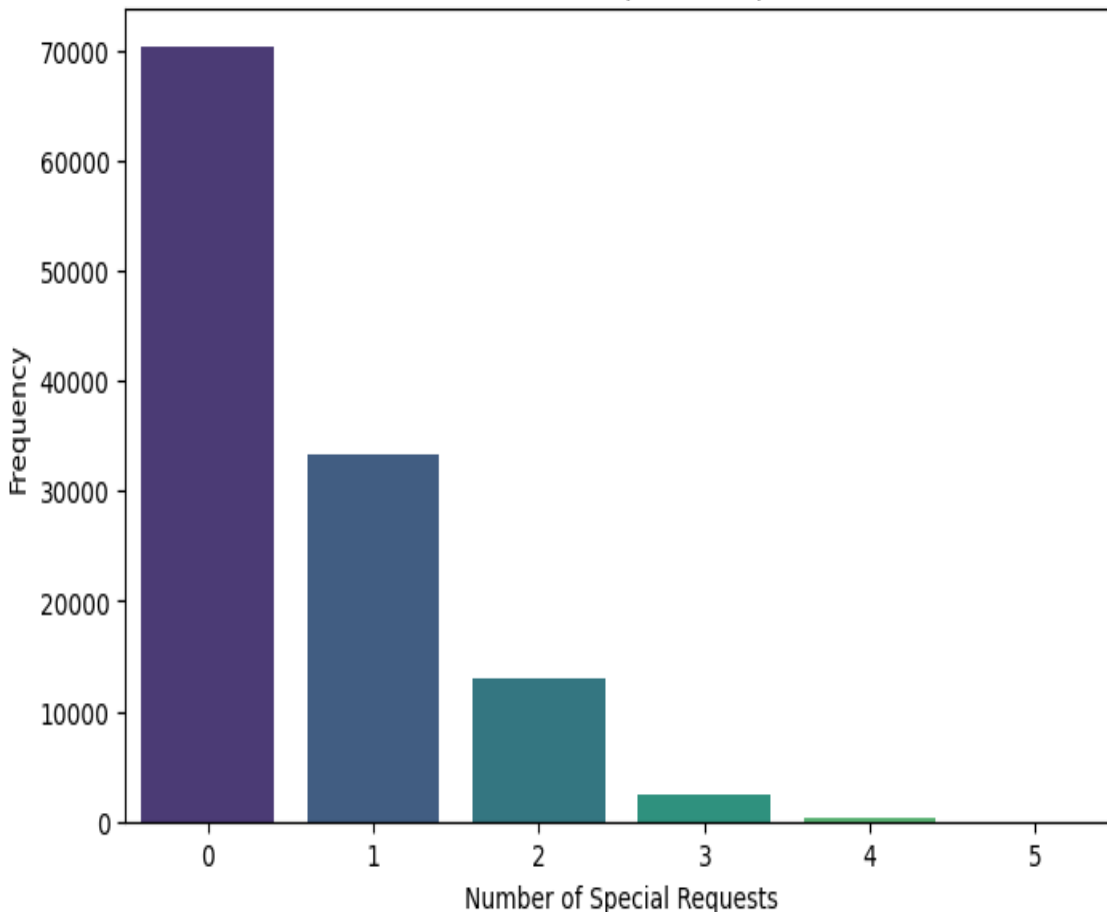
Distribution of Length of Stay



- **Distribution of length of stay:** This visualization gives you insights into the distribution of the length of stay, combining weekend and week nights.
- **Observation:** The hist-plot with bins and KDE suggests that the length of stay is most densely concentrated in 10 nights, indicated by the peak in the KDE curve, while the bins provide discrete counts within specific intervals

➡ Second problem as shown below

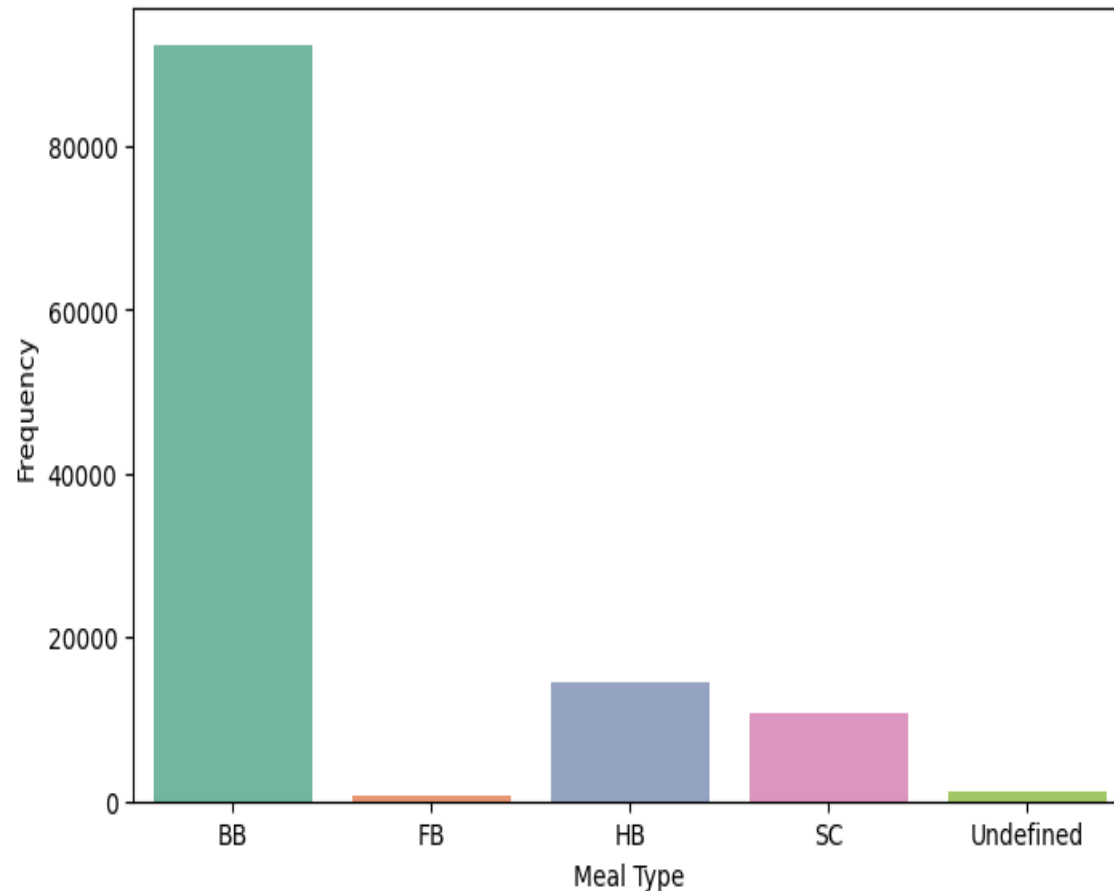
Distribution of Special Requests



- ▶ **Distribution of special request:** Second plot shows the distribution of the total number of special requests made by guests
- ▶ **Observation:** Frequency is high for Zero special requests which allows for a more streamlined and efficient service delivery, as the hotel staff can focus on consistently meeting standard expectations. This, in turn, contributes to a smoother and more hassle-free experience for guests, minimizing potential delays or errors associated with fulfilling specific requests.

➡ Third problem as shown below

Meal Preferences of Guests

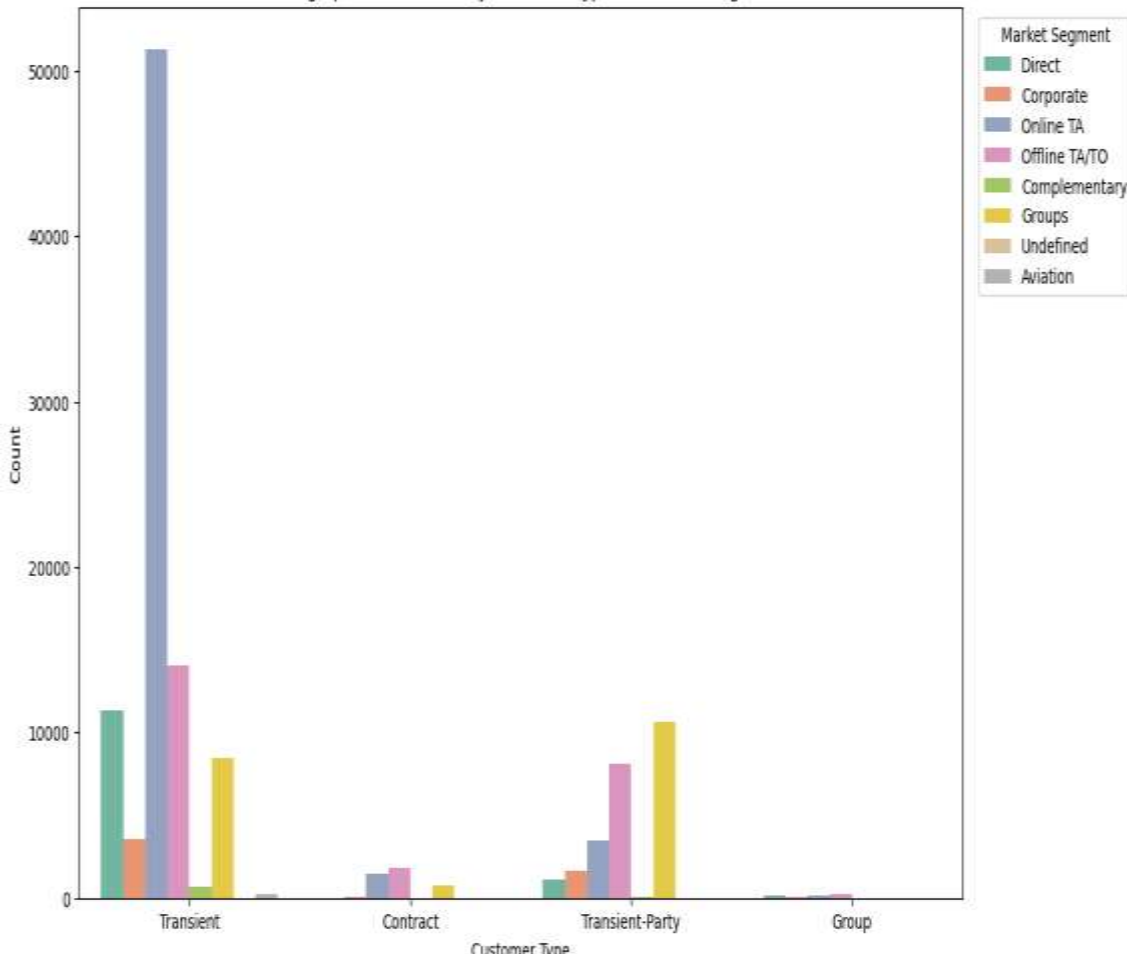


- ▶ **Meal preferences:** Count-plot displays the distribution of meal preferences (BB, FB, SC, HB) among guests.
- ▶ **Observation:** The frequency is highly reserved for Bread and Breakfast(BB) Meal Type. If we can provide special discount or offers on BB and Half-Board(HB) and Self-Catering(SC), we can enhance the customer experience

## Improving Marketing Efforts: This problem statement include few sub problems

➡ First problem as shown below

Demographic Distribution by Customer Type and Market Segment

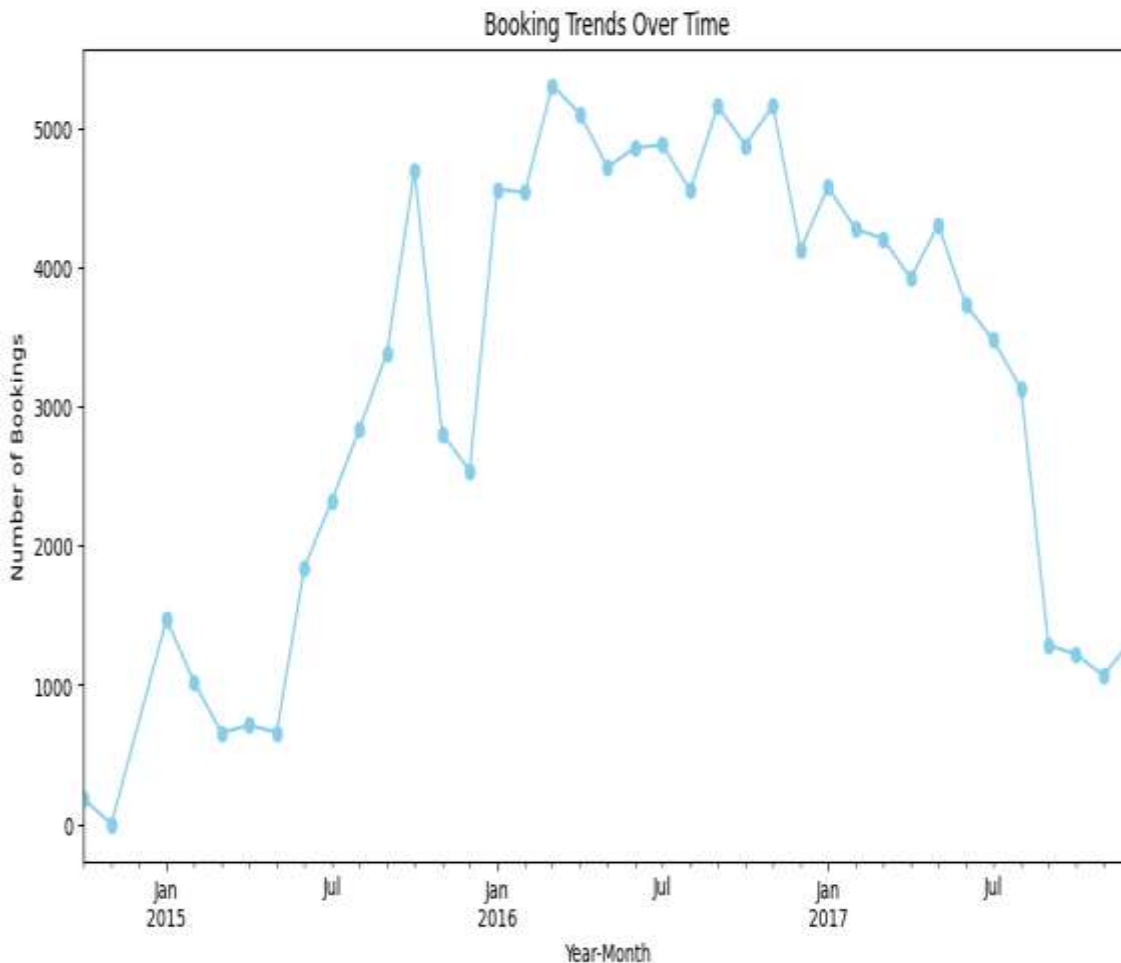


► **Demographic Distribution:** Stacked Bar Chart

► **Observation:** This chart reveals the distribution of market segments within each customer type, helping identify which segments are more prevalent among different types of guests. Transient customer type has highest distribution in Online Travel Agency(TA). Offering exclusive deals, loyalty programs, or bundled packages can attract and retain Transient customers, improving the effectiveness of marketing campaigns.



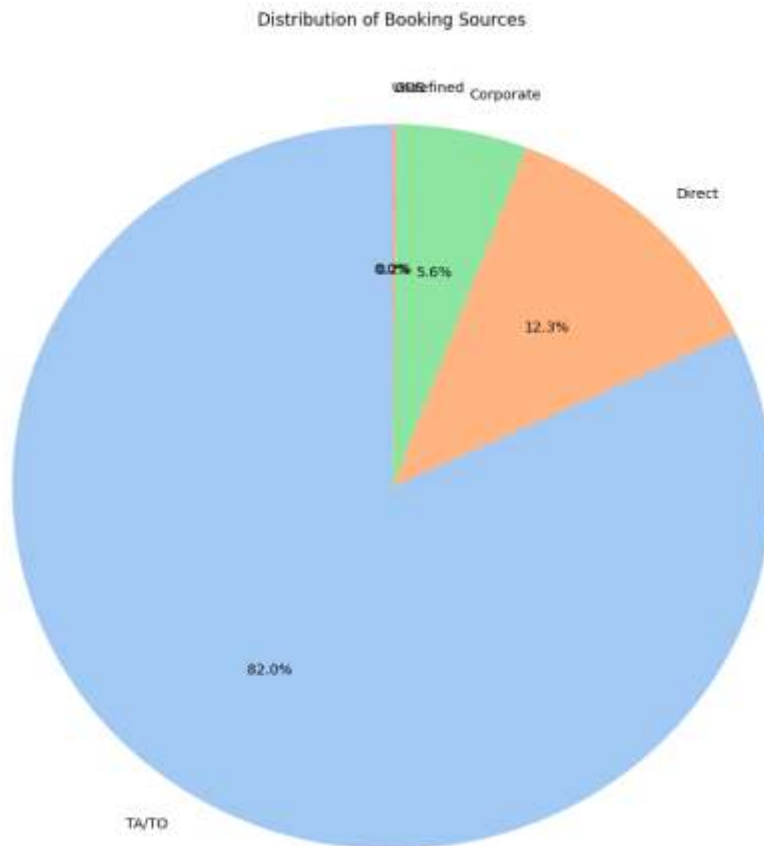
➡ Second problem as shown below



► **Booking Trends Over Time:** Line Chart

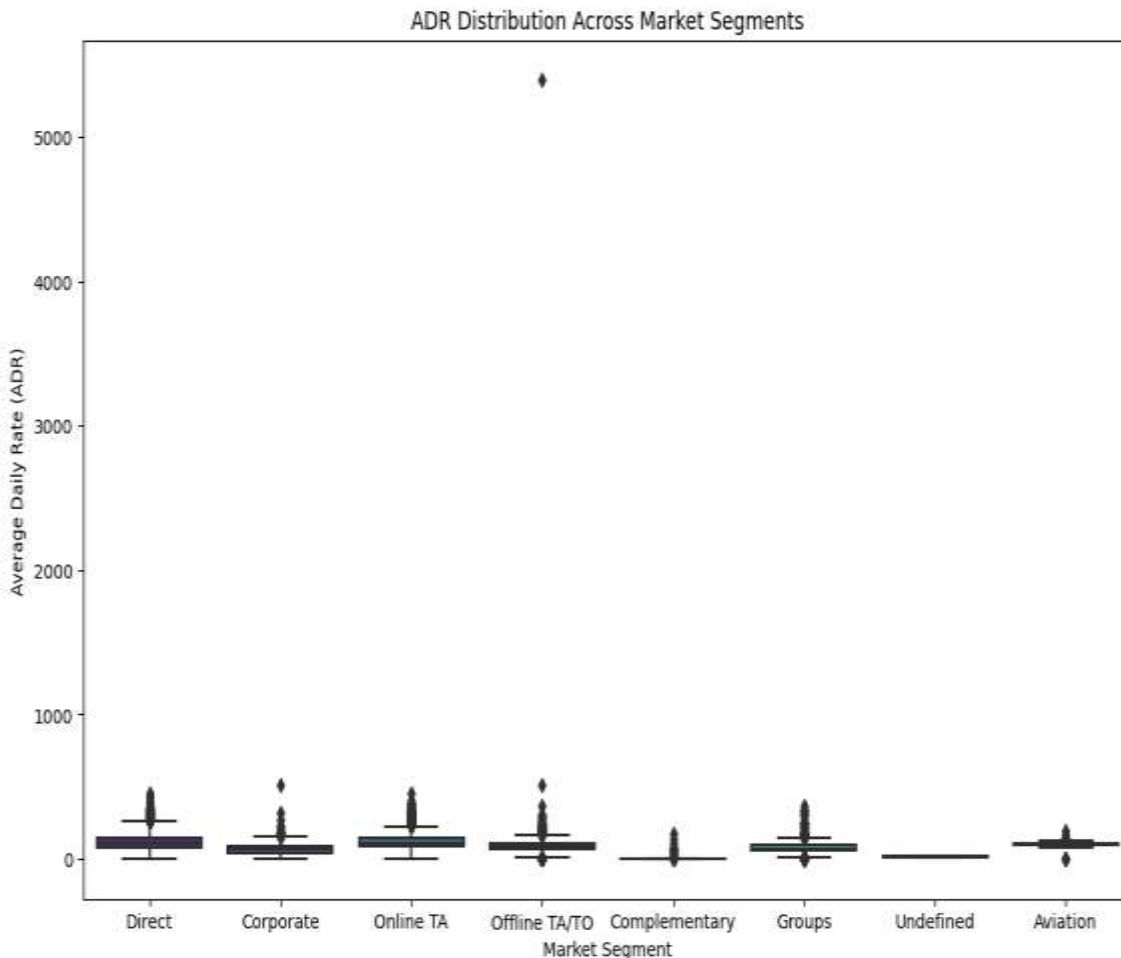
► **Observation:** This line chart illustrates booking trends over time, helping identify peak periods and seasonality for targeted marketing campaigns. If we see 3 years of line chart, the bookings ratio stable and increased from January 2016 to May 2017

➡ Third problem as shown below



- ▶ **Source of Bookings:** Pie Chart
- ▶ **Observation:** A pie chart depicting the distribution of booking sources helps allocate marketing resources effectively, emphasizing channels that contribute most to bookings. Travel Agency/Travel Office (TA/TO) contribute 82% of highest distribution of bookings

➡ Forth problem as shown below



► **Market Segment Analysis:** Box Plot

► **Observation:** A Box plot depicting the distribution of Average Daily Rate (ADR) helps allocate marketing resources effectively for various Booking market segments, emphasizing channels that contribute the ADR. As we can see the ADR is stable and high for Direct, Corporate, Online TA, Offline TA/TO and groups,

## Optimizing inventory Management and Forecasting Demand: These problem statements include few subproblems

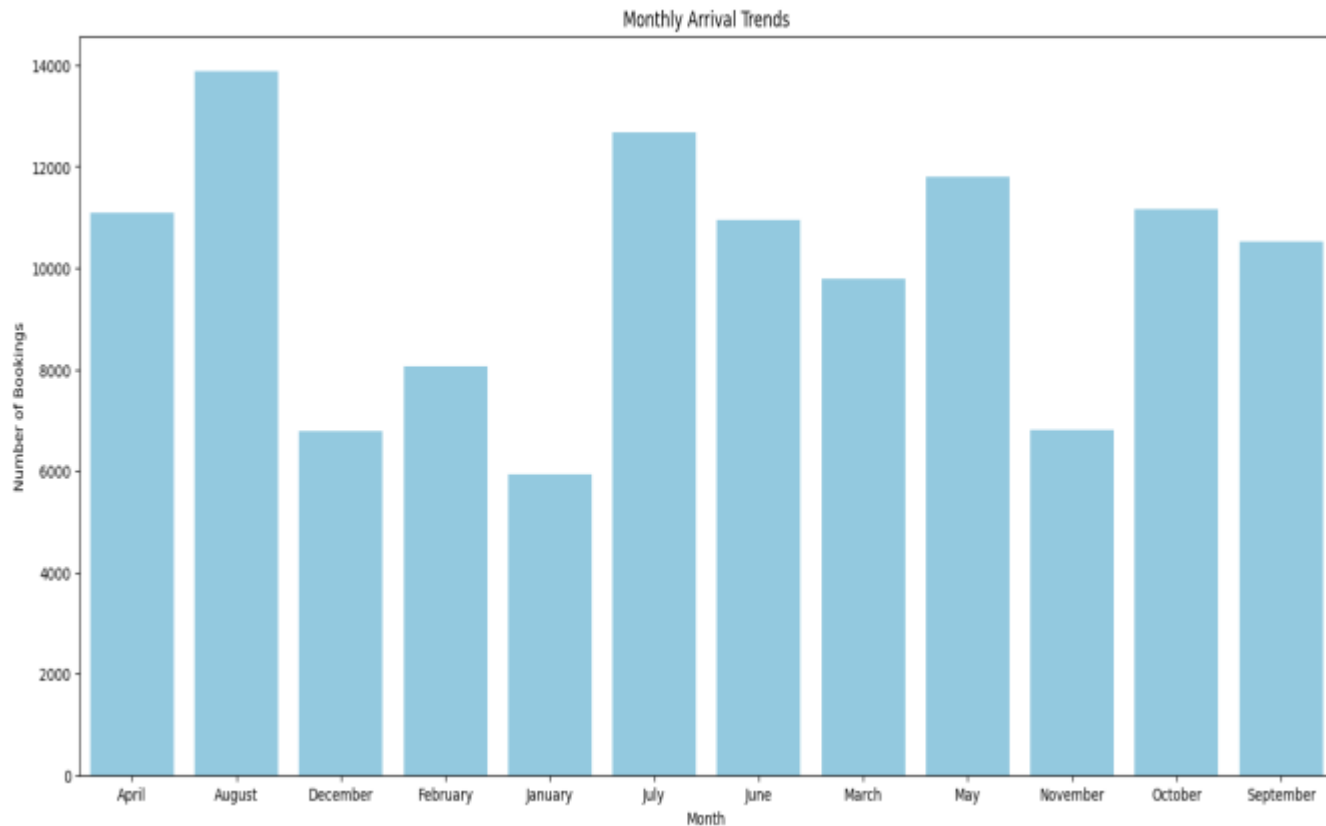
➡ First Problem as shown below

**Observation:** August is the first highest and July is the second highest booked months. Least booked months are January, November and December

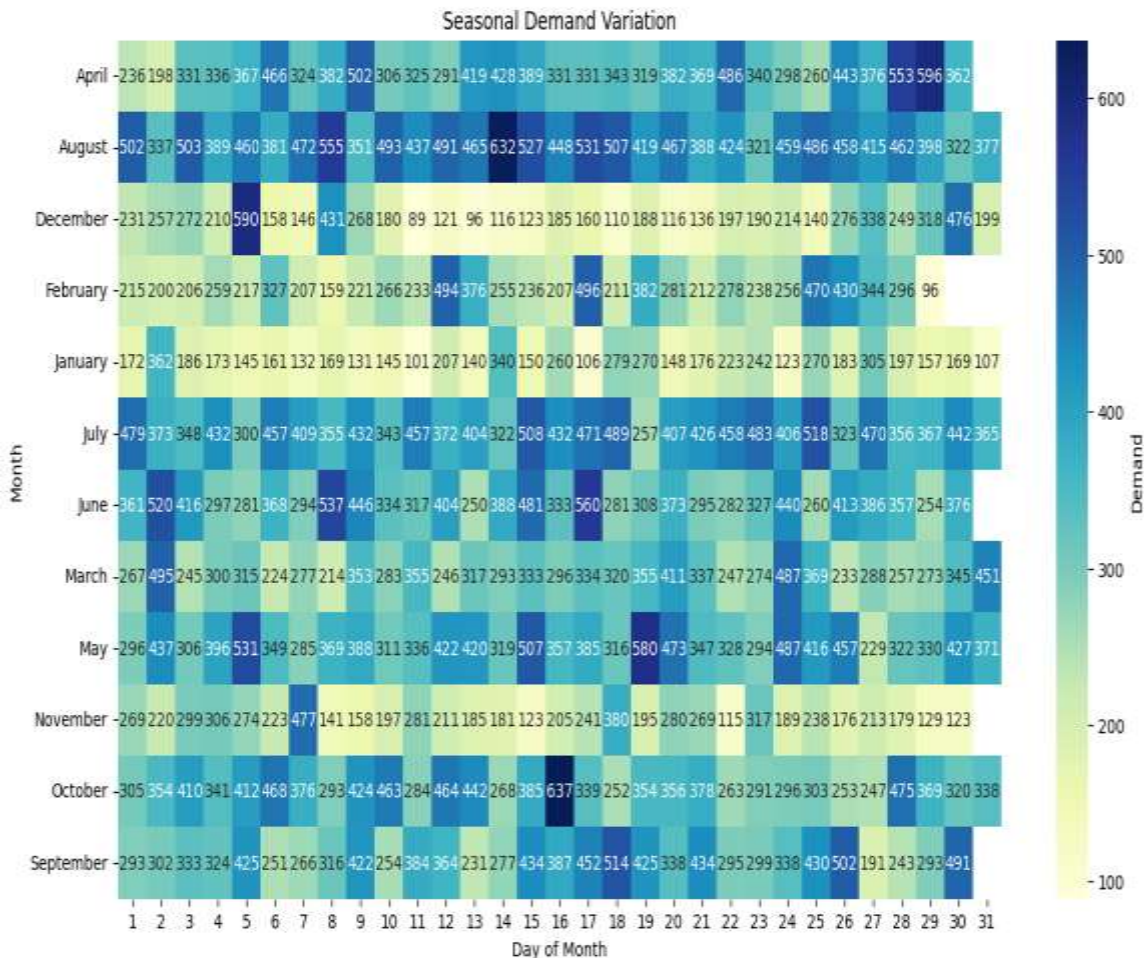
Number of Bookings for Each Month:

August	13877
July	12661
May	11791
October	11160
April	11089
June	10939
September	10508
March	9794
February	8068
November	6794
December	6780
January	5929

Name: arrival\_date\_month, dtype: int64



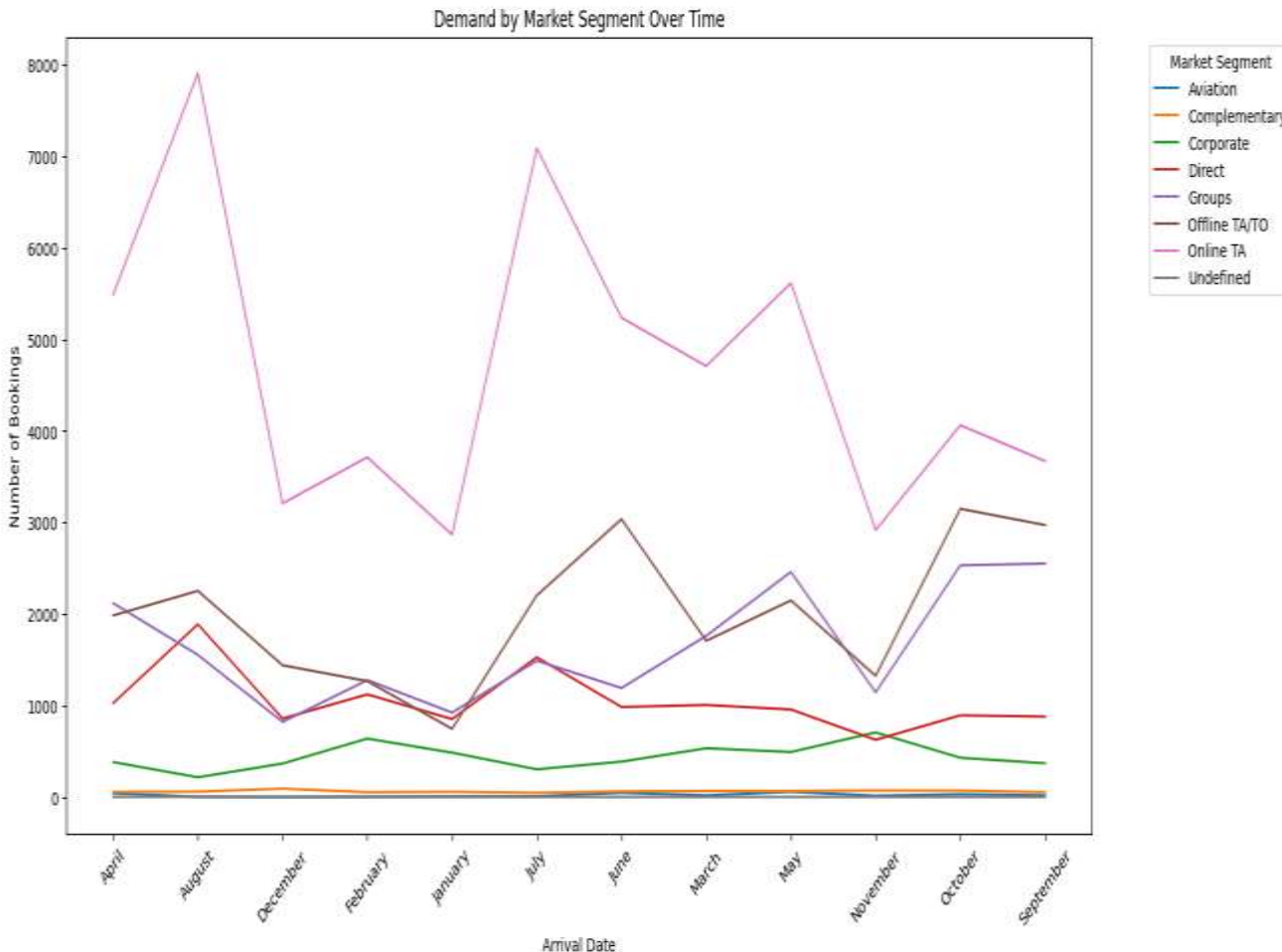
## ➔ Second Problem as shown below



- **Seasonal demand variation:** heat-map to identify the demand day of month wise
- **Observation:** There are few months like May, June, July, august, September and October are the month of In-demand seasons

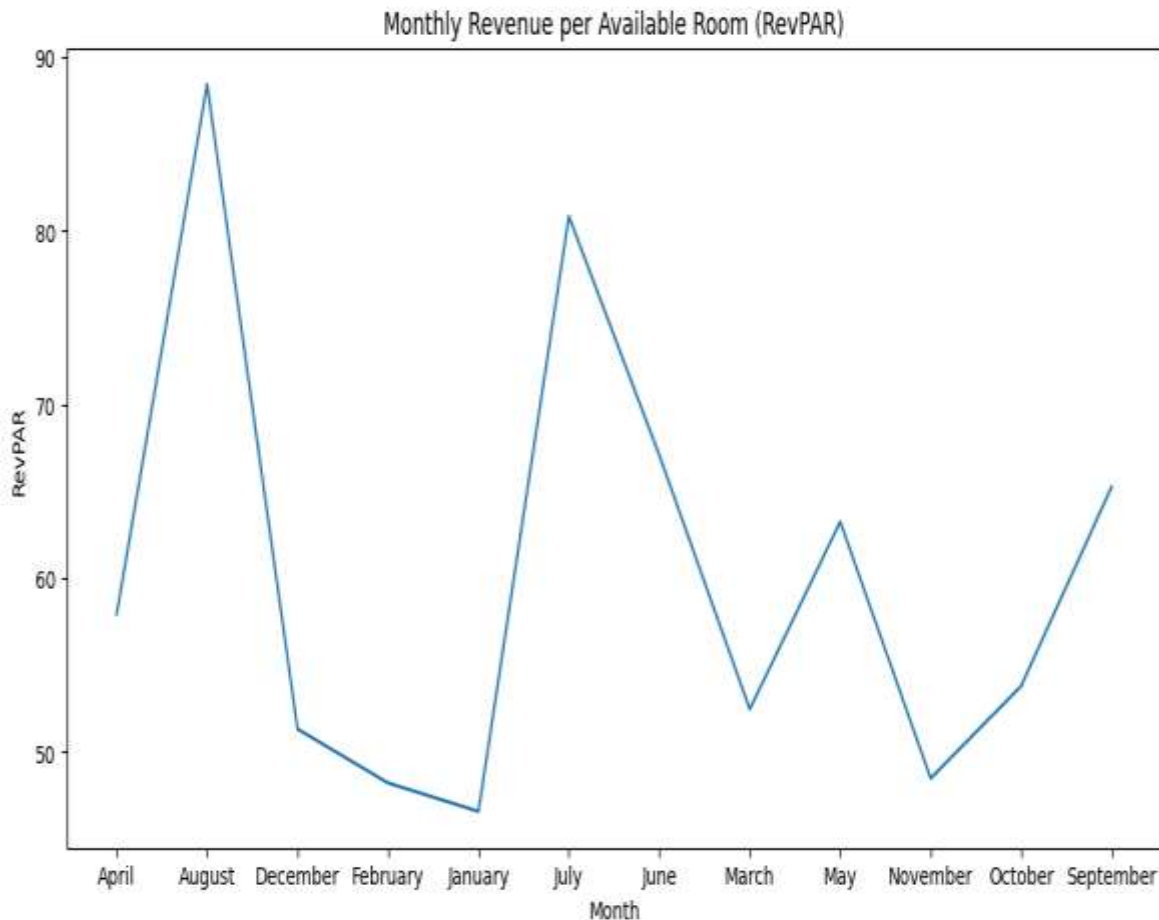


## ➡ Third Problem as shown below



**Observation :** Demand by market segment over Time: line chart that shows the total number of bookings for each market segment across different periods. Online Travel Agency has the potential of acquiring highest number bookings mostly in the month of August and July

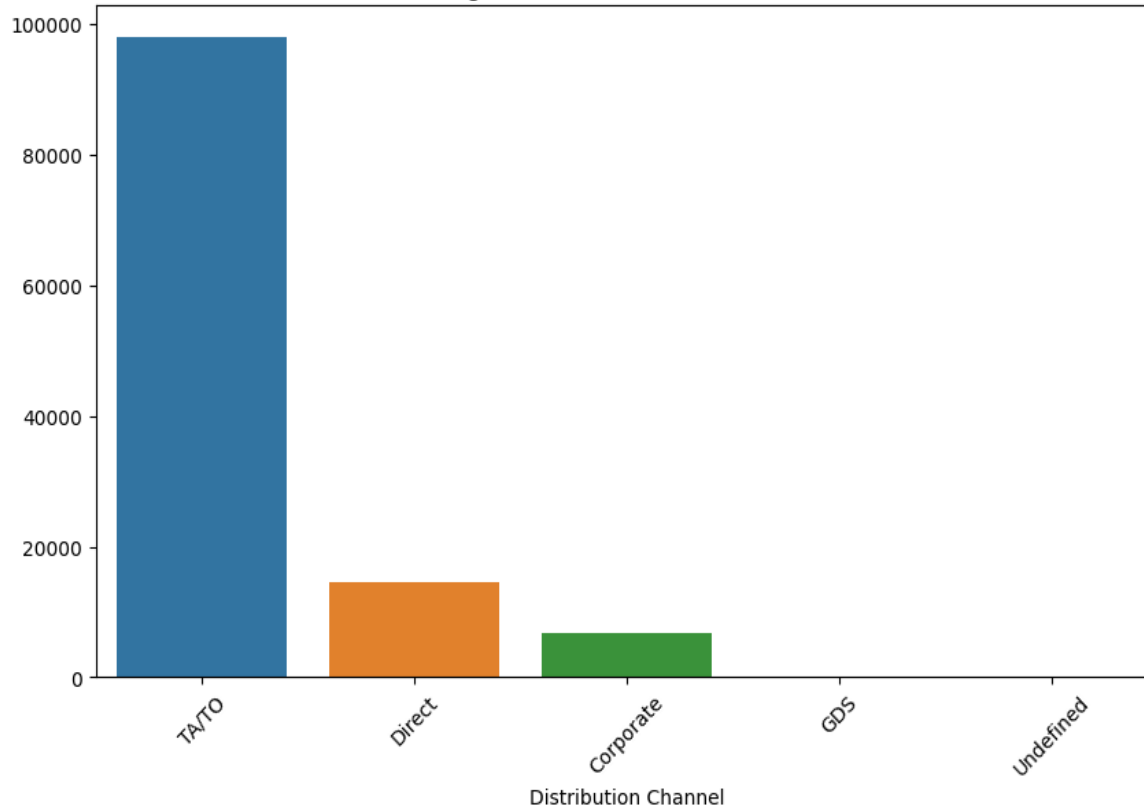
➡ Forth Problem as shown below



- **Observation:** So the Monthly Revenue per Available Room is derived by using Line plot which helps to Observe that the August and July are considered to be the high revenue months for per room available for booking

➡ Fifth Problem as shown below

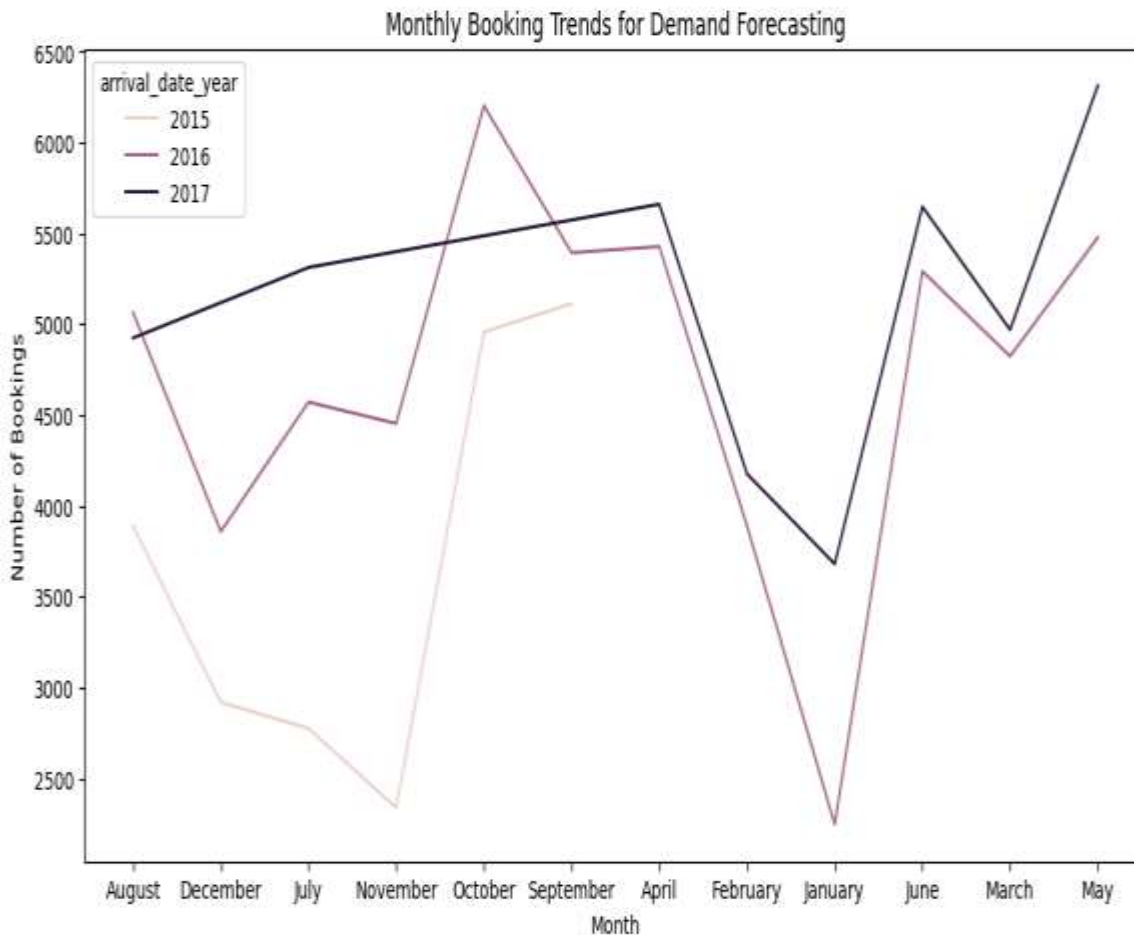
Booking Channel Contribution to Demand



- ▶ **Booking channel contribution to Demand:** Bar plot
- ▶ **Observation:** So the Demand for hotel is through TA/TO (Travel Agency and Travel Office) is peak

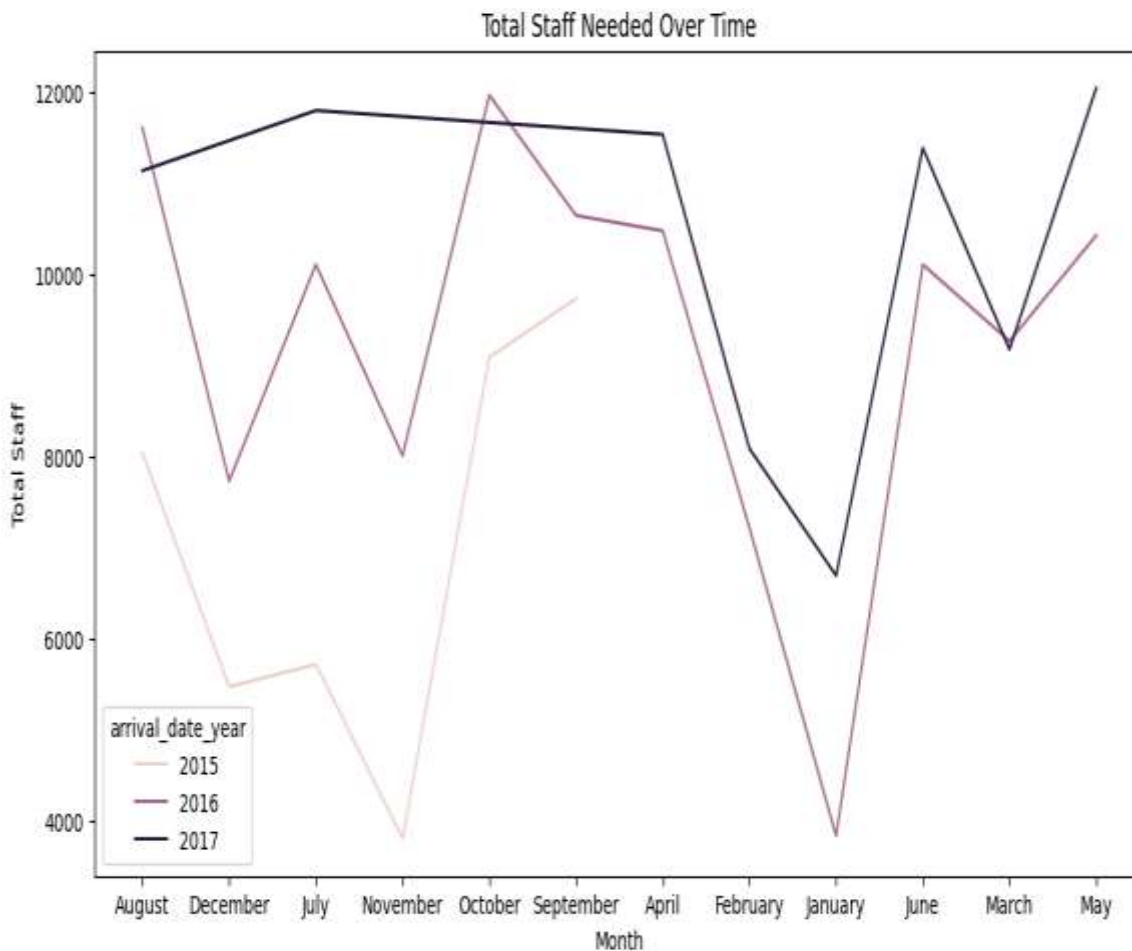
## Reducing Operational Cost: It includes sub problems

### First problem as shown below



- ▶ **Month wise booking trend:** Visualize monthly booking trends for demand forecasting
- ▶ **Observation:** This visualization helps in understanding the monthly booking trends, aiding in demand forecasting and resource planning. The demand for bookings increased in May in 2017 compared to last year 2016

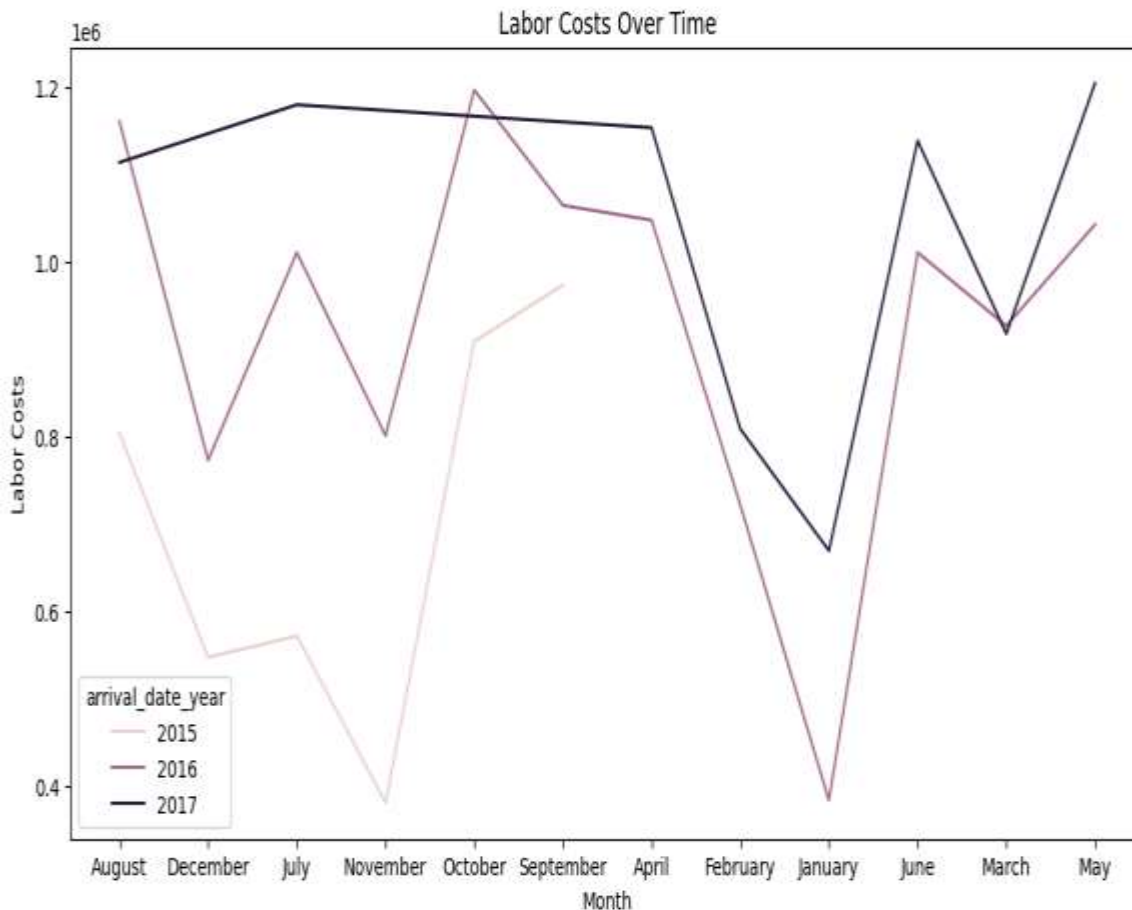
## ➡ Second problem



- **Staff need:** Visualizing the need of staff as per the booking for monthly basis
- **Observation:** For the year of 2017, there is an increase and the stable requirement of Staff need apart from the months January, February and March.

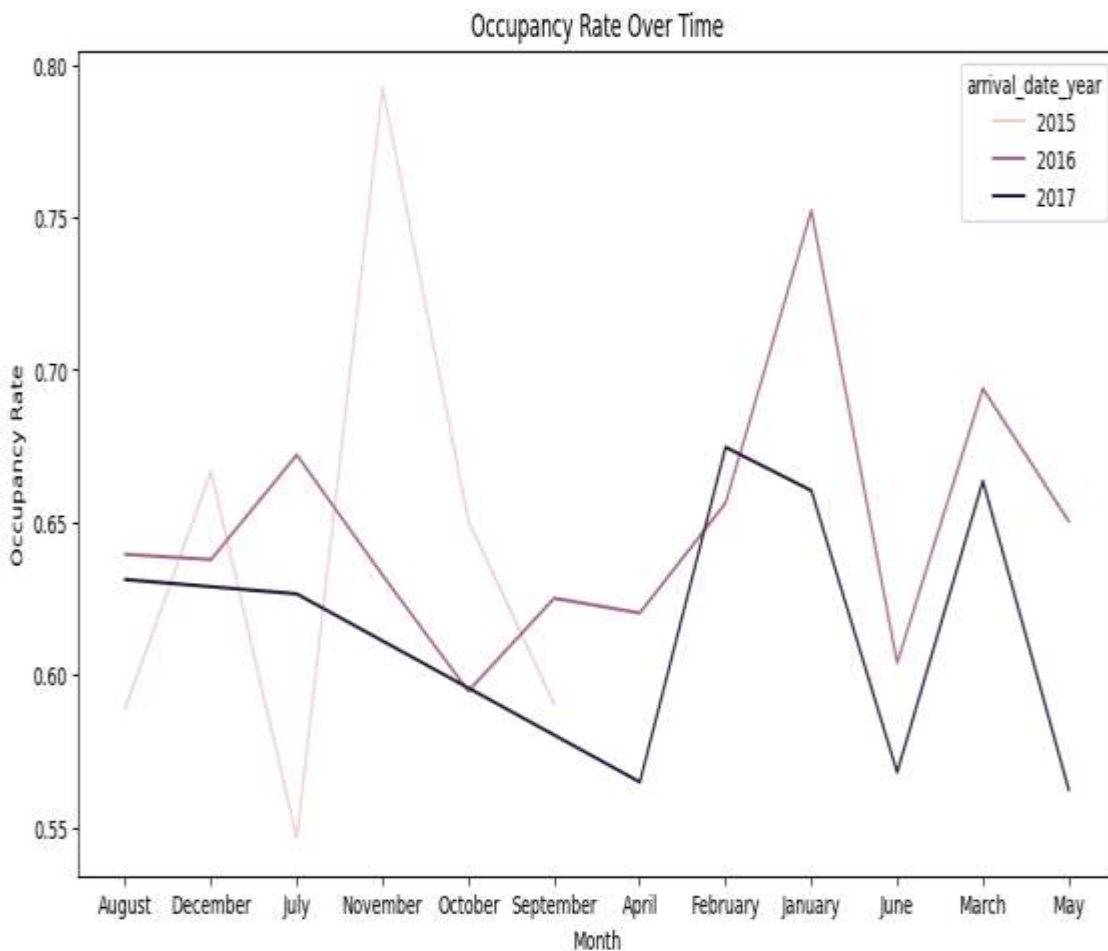


## ➡ Third problem



- **Observation:** For the year of 2017, there is an increase and the stability in the labor cost apart from the months January, February and March (Downfall of the labor cost)

## ➡ Forth problem

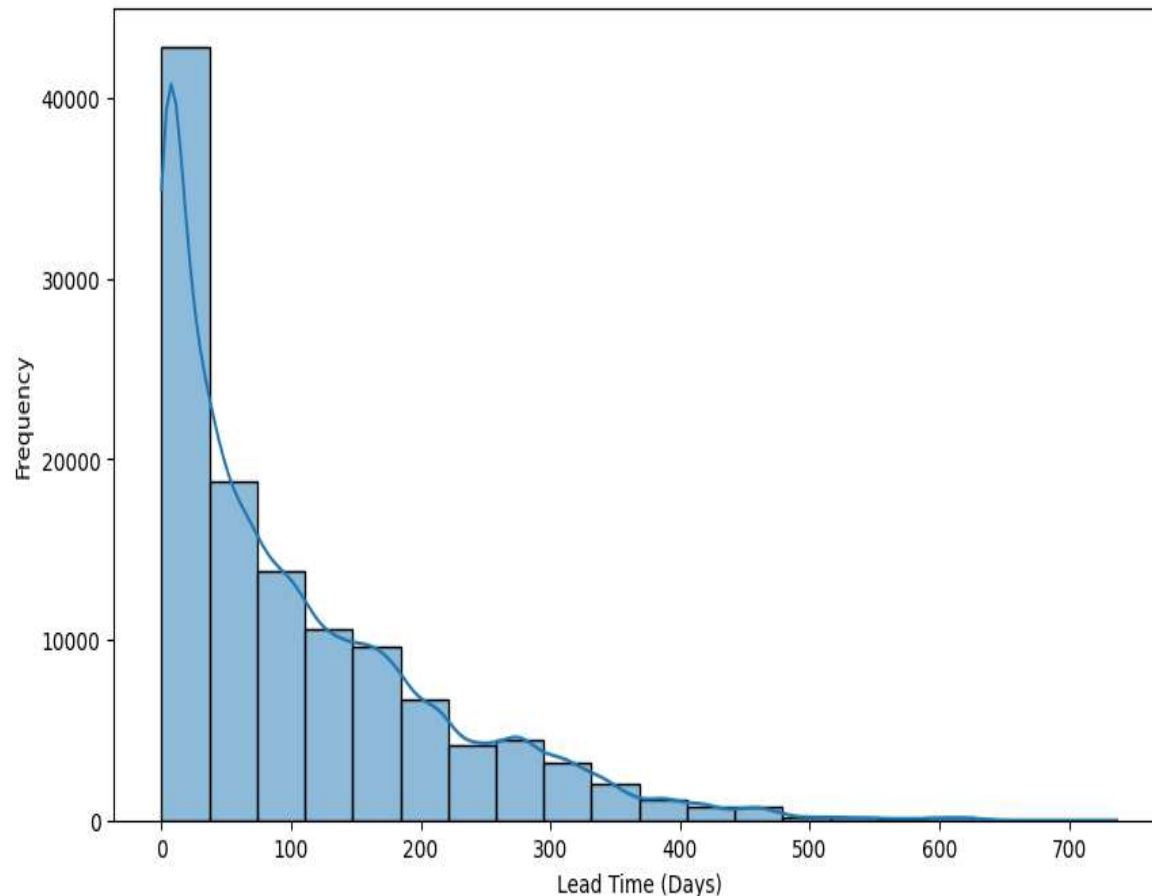


- **Observation:** Room occupancy rates are less in may, June ,April September and October for the year 2017

Enhancing operational efficiency: This includes many sub problems

➡ First sub problem as shown below

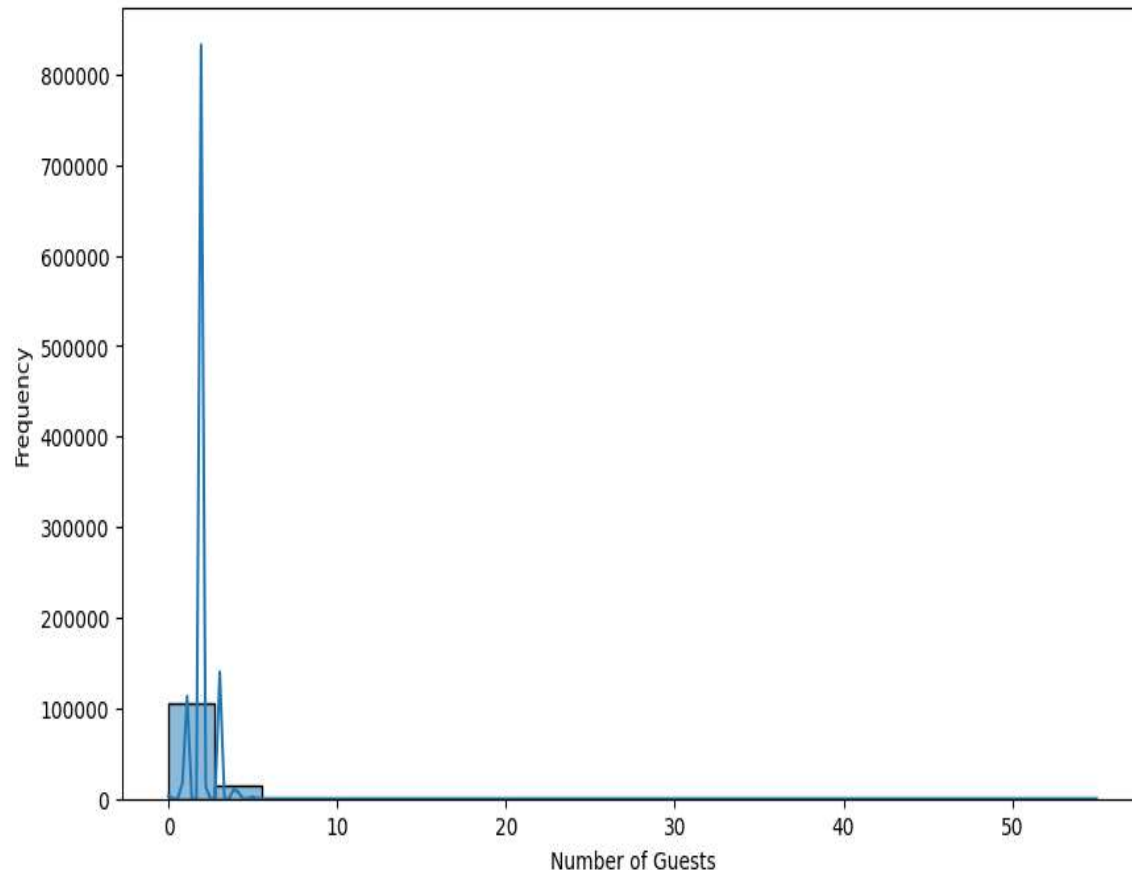
Distribution of Lead Time



- **Visualize the distribution of lead time** :time between booking and check-in
- **Observation:** This chart helps in understanding the lead time distribution, which can influence the efficiency of check-in/check-out processes. So the Leading time of days for the first 25 to 30 days are high

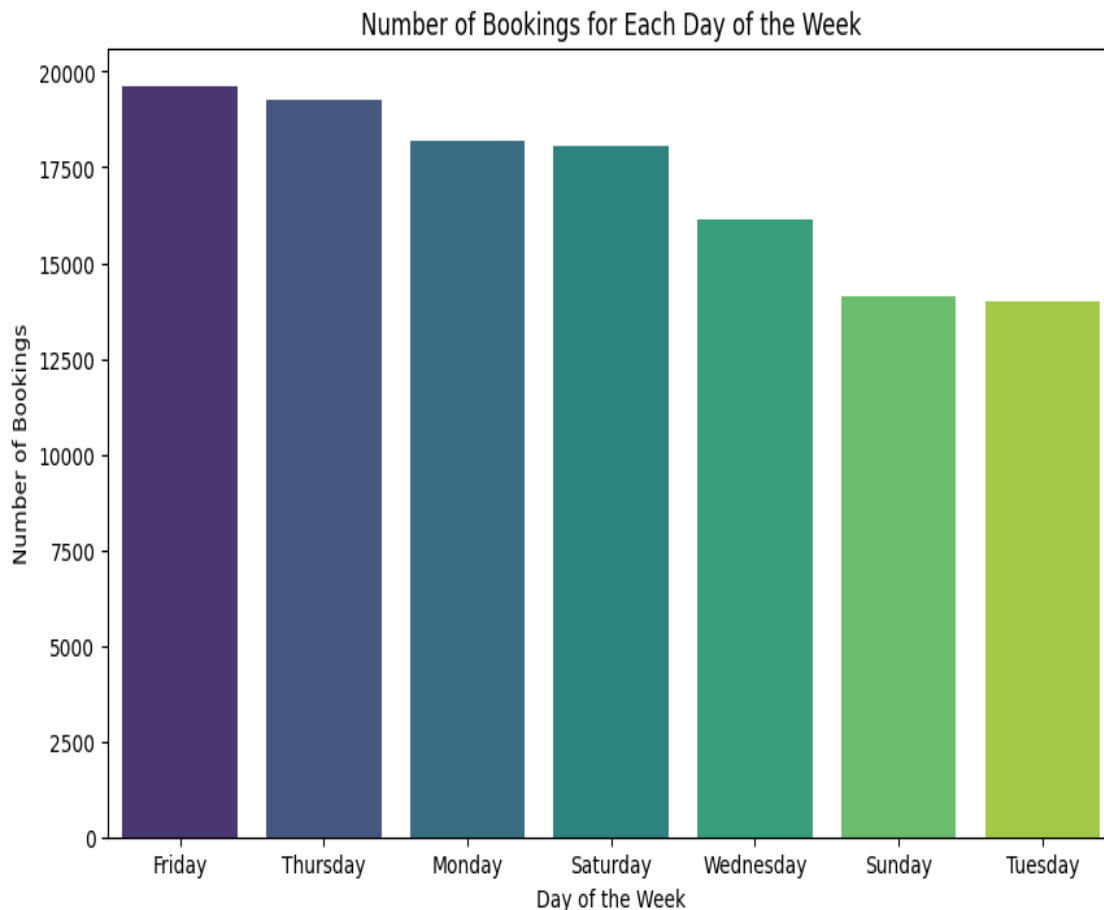
## → Second problem

Distribution of Total Guests



- ▶ **Resource Allocation:** Visualize the distribution of total guests (adults + children + babies)
- ▶ **Observation:** This visualization provides insights into the distribution of the total number of guests, helping in resource allocation planning. Average 5 to 6 guests are considered to be the most booked guests

## → Third problem



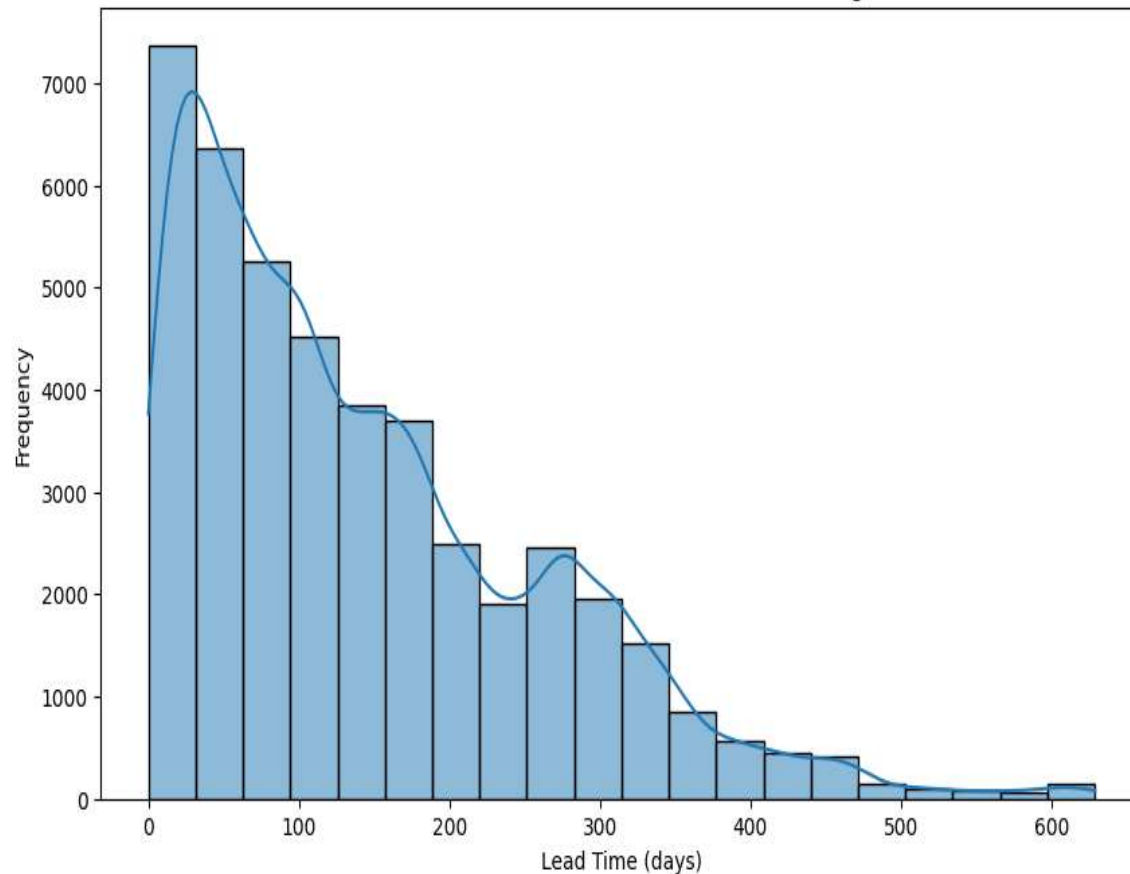
- **Busy Period Preparation:** Number of Bookings for each day of week
- **Observation:** Preparing for busy periods by understanding the distribution of bookings across different days of the week. Friday, Thursday is the peak booking day for a week



**Risk management: This include sub problems as shown below**

**➡ First problem**

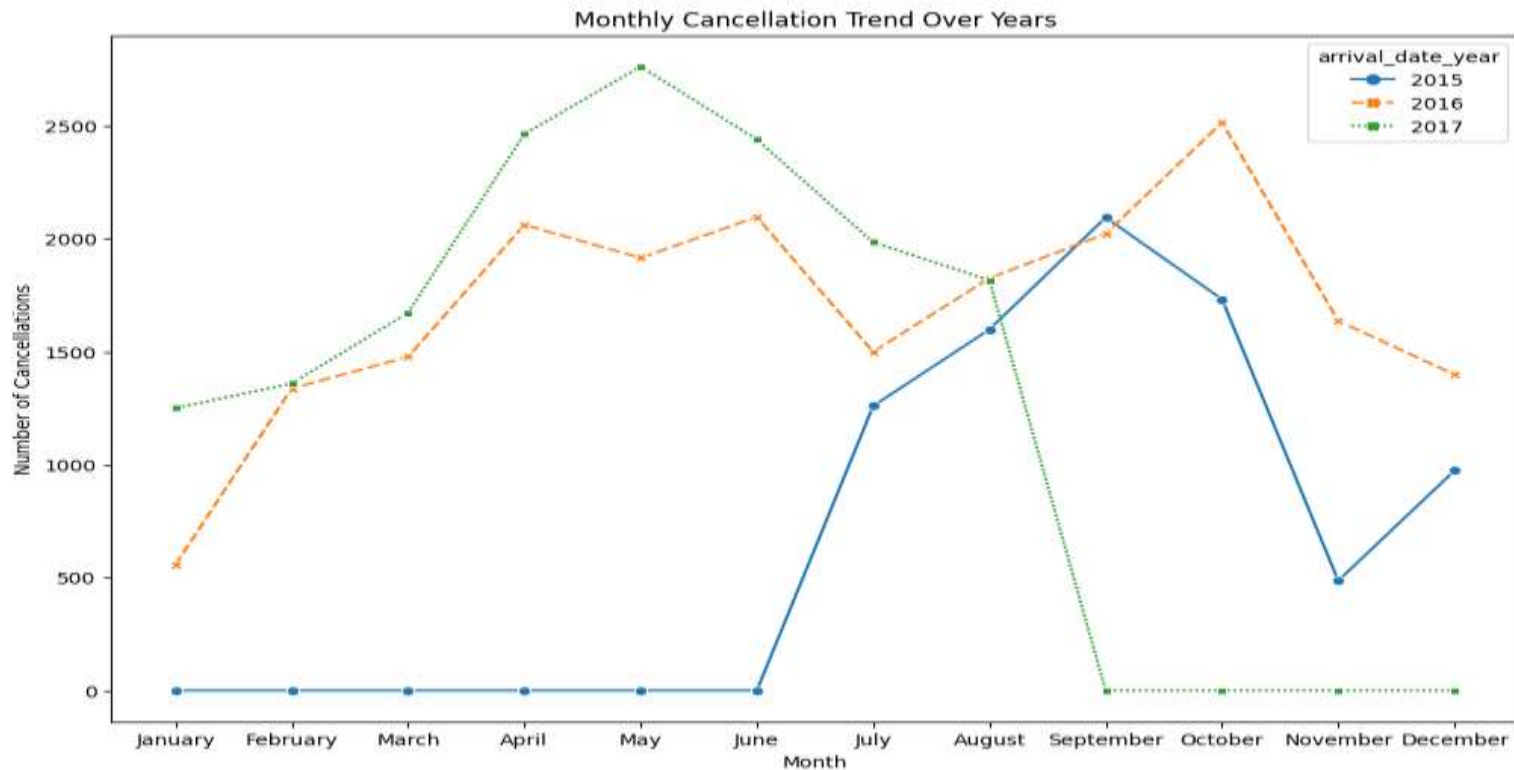
Lead Time Distribution for Canceled Bookings



- In the context of hotel bookings, a higher concentration of cancellations within the initial 25 to 30 days of lead time could indicate that guests are more likely to cancel their reservations shortly after making them. This insight is valuable for hotel management to consider when implementing cancellation policies, adjusting pricing strategies, or managing inventory to mitigate potential revenue losses associated with cancellations.

## ➡ Second problem

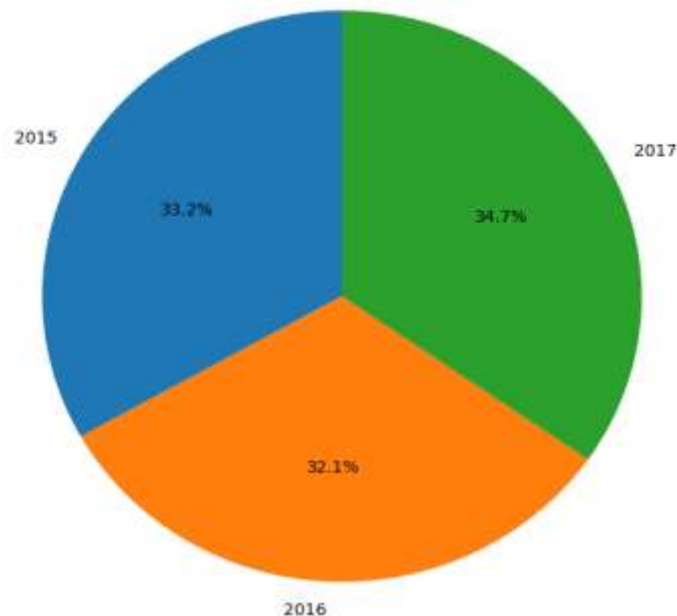
**Observation:** In the year 2017, the month June, April and May has a high number of cancellations and if we observe the market segment,



## → Third problem

**Observation:** In the year 2017, the month march, April and may has high number of cancellations through Online Travel Agency (TA) which can increase the risk for the revenue growth curve

Percentage of Bookings Canceled by Year

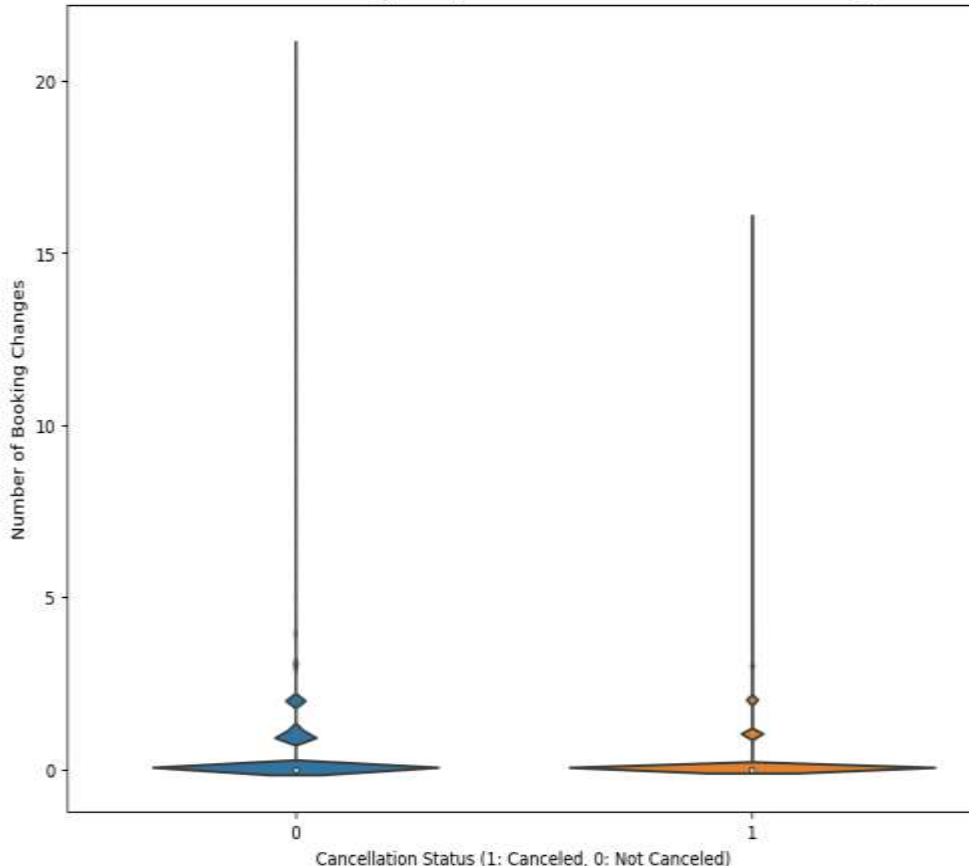


Pivot Table - total number of Cancellations :

	sum		
arrival_date_year	2015	2016	2017
market_segment			
Aviation	0	29	23
Complementary	15	47	35
Corporate	219	375	398
Direct	351	867	716
Groups	3843	4852	3402
Offline TA/TO	2067	4330	1914
Online TA	1645	9837	9257
Undefined	2	0	0

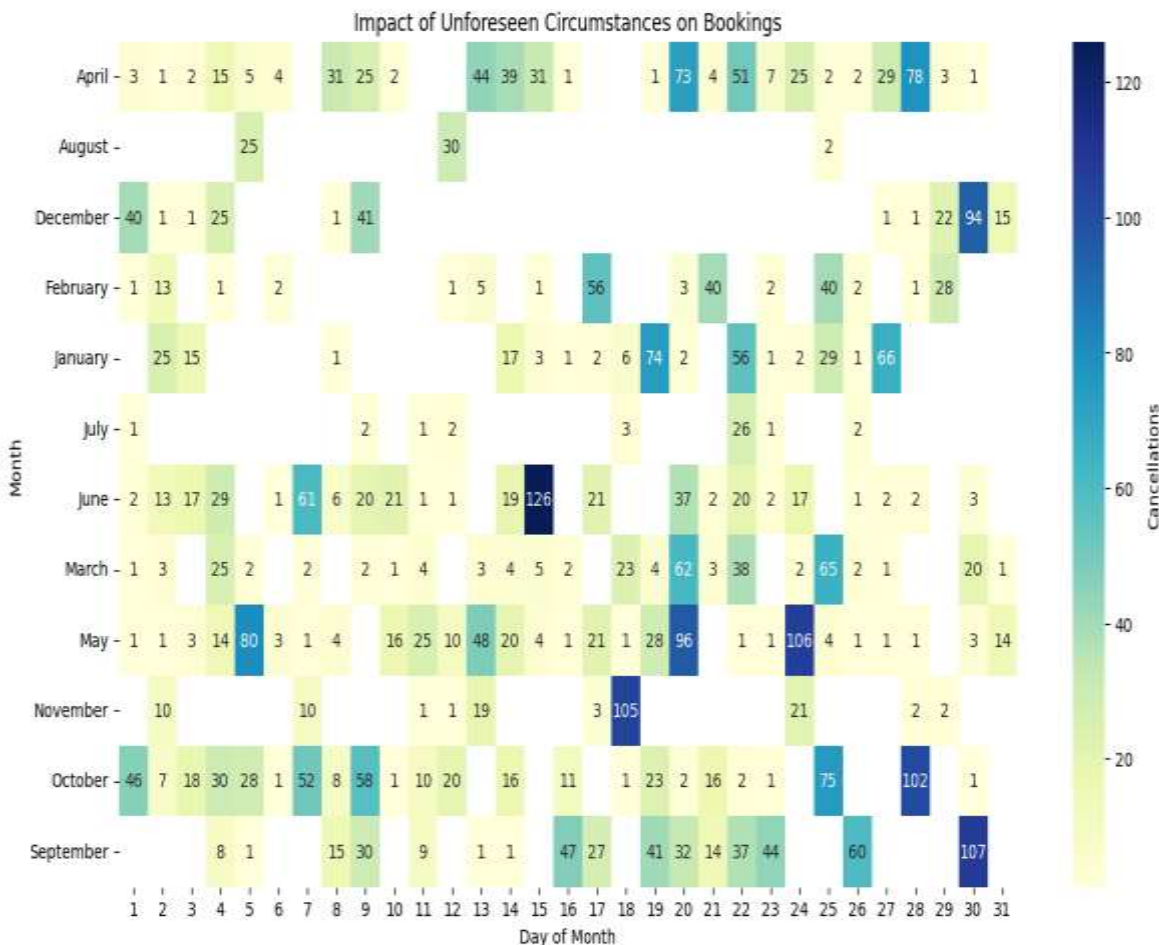
## ➡ FORTH problem

Distribution of Booking Changes for Canceled vs. Not Canceled Bookings



- ▶ **Booking Changes and Cancellation Correlation:** A violin-plot showing the correlation between the number of booking changes and the likelihood of cancellation can highlight potential correlations.
- ▶ **Observation:** Cancellation rate is increasing if there is any booking changes made at the time of reservation thus it is the risk of cancellation of booking and this will help in identifying patterns where frequent changes may lead to increased cancellations.

## ➔ Fifth problem



- **Impact of Unforeseen Circumstances on Bookings:** Heat-map helps to gain insights day of month wise bookings
- **Observation:** There are few days of month May, November, June, October, September have seen some Unforeseen circumstances on booking which can increase the risk for bookings

## Conclusion:

### Insights that help in creating a positive business impact

- ▶ Optimizing pricing strategies involves implementing dynamic pricing during peak seasons like July and August, introducing tailored packages for shorter stays (2 to 15 nights) to attract a broader audience, and implementing segmented pricing, capitalizing on the higher Average Daily Rate (ADR) for Transient customers, particularly Transient-party, which exhibits a high requirement for car parking spaces, thereby maximizing revenue and refining overall pricing dynamics.
- ▶ Concentrating on the most densely populated stay duration of 10 nights, hotels can tailor services, promotions, and room arrangements to align with this preference, contributing to a more personalized and satisfying guest experience. Offering special discounts on the preferred Bread and Breakfast (BB) and Half-Board (HB) meal types identified in the analysis further enhances customer satisfaction by catering to specific culinary preferences and enriching the overall stay.
- ▶ Leveraging insights from market segment distribution, booking trends, and source contributions, strategic marketing efforts can prioritize exclusive deals for Transient customers on Online Travel Agencies, align campaigns with stable and increasing booking trends, strategically allocate resources to prevalent channels like Travel Agency/Travel Office (TA/TO), and optimize pricing strategies for segments with high and stable Average Daily Rates (ADR). This holistic approach ensures a targeted, timely, and resource-efficient marketing strategy, enhancing overall effectiveness and maximizing customer engagement and revenue potential.
- ▶ The observation of July and August being the highest booking months not only aids in strategic staffing but also guides revenue optimization efforts. By aligning promotions or packages to attract more bookings, maximizing revenue per room and staffing with peak demand, hotels can enhance guest experiences and capture the revenue potential associated with these high-demand periods. Hotels can use booking data to optimize room inventory management. Specifically August and July are the months to maximize the revenue per available room. This involves adjusting room availability based on historical booking trends, minimizing overbooking or underutilization of rooms, and maximizing revenue per available room.

## Insights that help in creating a positive business impact

- By analyzing lead time distribution, total guest numbers, and booking patterns across different days, the hotel gains actionable insights to enhance operational efficiency. The identification of high lead times informs streamlined check-in/check-out processes, accommodating guests who plan well in advance. Efficient resource allocation, tailored to the average 5 to 6 guests per booking, optimizes staffing and service provisioning.
- Additionally, recognizing Friday and Thursday as peak booking days enables strategic staffing, preparing the hotel for increased demand on these key days. Overall, these insights empower the hotel to align operations with guest behaviors, fostering an enhanced and efficient guest experience.
- The identified patterns of increased cancellations during specific months, especially in May, November, June, October, and September, provide a foundation for proactive risk management. By recognizing the correlation between booking changes and higher cancellation rates, hotels can implement stringent policies or dynamic pricing strategies to minimize cancellations. Additionally, focusing on the market segment, particularly Online Travel Agencies, allows for targeted risk mitigation strategies, such as refining partnerships or optimizing promotional activities. Understanding the concentration of cancellations within the initial 25 to 30 days of lead time enables the establishment of tailored policies, helping hotels to strategically navigate potential revenue risks and enhance overall financial resilience.
- Harnessing data analytics insights, such as identifying peak cancellation periods, adapting to market trends, and understanding booking patterns, equips hotels with a competitive advantage. By staying ahead of trends and efficiently managing operational processes, hotels can enhance customer satisfaction and loyalty. The ability to adapt quickly to changes, fueled by data-driven Decision-making, positions hotels to meet evolving customer preferences. Continuous improvement in services, informed by detailed analytics, enables hotels to offer tailored experiences, creating a distinct competitive edge in the dynamic hospitality industry.





**Thank you**